

日 本 国 特 許 庁  
JAPAN PATENT OFFICE

PCT/JP 2004/006065

27. 4. 2004

Rec'd PCT/PTO 21 JUL 2005

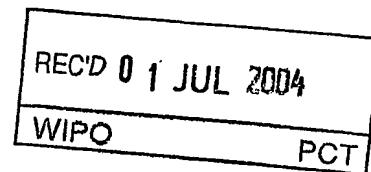
別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office.

出 願 年 月 日  
Date of Application: 2003年 5月21日

出 願 番 号  
Application Number: 特願2003-143043  
[ST. 10/C]: [JP 2003-143043]

出 願 人  
Applicant(s): 松下電器産業株式会社

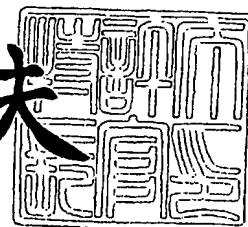


PRIORITY DOCUMENT  
SUBMITTED OR TRANSMITTED IN  
COMPLIANCE WITH  
RULE 17.1(a) OR (b)

2004年 6月 3日

特許庁長官  
Commissioner,  
Japan Patent Office

今井康夫



Best Available Copy

出証番号 出証特2004-3047751

【書類名】 特許願  
【整理番号】 2931030151  
【提出日】 平成15年 5月21日  
【あて先】 特許庁長官殿  
【国際特許分類】 G10K 15/22  
【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 西崎 誠

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 小沼 知浩

【発明者】

【住所又は居所】 大阪府門真市大字門真 1 0 0 6 番地 松下電器産業株式会社内

【氏名】 遠藤 充

【特許出願人】

【識別番号】 000005821

【氏名又は名称】 松下電器産業株式会社

【代理人】

【識別番号】 100097445

【弁理士】

【氏名又は名称】 岩橋 文雄

【選任した代理人】

【識別番号】 100103355

【弁理士】

【氏名又は名称】 坂口 智康

【選任した代理人】

【識別番号】 100109667

【弁理士】

【氏名又は名称】 内藤 浩樹

【手数料の表示】

【予納台帳番号】 011305

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9809938

BEST AVAILABLE COPY

【書類名】 明細書

【発明の名称】 音声応答装置及び音声応答方法

【特許請求の範囲】

【請求項 1】 文字情報及び画像情報を表示する応答テキスト出力部と、音声情報を音声として出力する音声出力部と、前記文字情報が出力される時刻と前記音声情報が出力される時刻とに遅延時間を設定する応答出力管理部とを含む音声応答装置。

【請求項 2】 遅延時間は、ユーザとの対応を行うキャラクタであるエージェントの応答テキスト出力部への表示の有無をパラメータとして決定される値である請求項 1 記載の音声応答装置。

【請求項 3】 遅延時間は、応答テキスト出力部に表示される文字情報とエージェントとの距離をパラメータとして決定される値である請求項 2 記載の音声応答装置。

【請求項 4】 遅延時間は、応答テキスト出力部に表示される文字情報の表示位置をパラメータとして決定される値である請求項 1 ないし 3 のいずれか記載の音声応答装置。

【請求項 5】 遅延時間は、ユーザ特徴を表すプロフィール情報をパラメータとして決定される値である請求項 1 ないし 4 のいずれか記載の音声応答装置。

【請求項 6】 遅延時間は、操作回数、利用時間を含むユーザの利用状況をパラメータとして決定される値である請求項 1 ないし 5 のいずれか記載の音声応答装置。

【請求項 7】 ユーザの視線位置を計測する視線位置計測部を更に含み、遅延時間は、前記ユーザの視線位置をパラメータとして決定される値である請求項 1 ないし 6 のいずれか記載の音声応答装置。

【請求項 8】 ユーザの呼吸の状態を測定する生体情報測定部を更に含み、遅延時間は、前記ユーザの呼吸の状態をパラメータとして決定される値である請求項 1 ないし 7 のいずれか記載の音声応答装置。

【請求項 9】 応答対象である文字情報は、点滅、応答対象でない文字情報よりも大きく表示又は応答対象でない文字情報とは異なる色で表示のうちのいずれ

か、あるいはそれらを組み合わせて表示される請求項 1 ないし 8 のいずれか記載の音声応答装置。

【請求項 10】 応答テキスト出力部は、文字情報を表示する場合に、ユーザの注意を喚起する報知音を出力する請求項 1 ないし 9 のいずれか記載の音声応答装置。

【請求項 11】 文字情報が出力される時刻と音声情報が出力される時刻とに遅延時間を設定する第 1 のステップと、文字情報を表示する第 2 のステップと、前記遅延時間の経過後に、音声情報を音声として出力する第 3 のステップとを含む音声応答方法。

【請求項 12】 文字情報が出力される時刻と音声情報が出力される時刻とに遅延時間を設定する第 1 のステップと、文字情報を表示する第 2 のステップと、前記遅延時間の経過後に、音声情報を音声として出力する第 3 のステップとを含む音声応答方法を実行するプログラム。

#### 【発明の詳細な説明】

##### 【0001】

##### 【産業の利用分野】

本発明は、ユーザの発声した内容に対して音声を用いて応答する音声応答装置に関するものである。

##### 【0002】

##### 【従来技術】

従来の音声応答装置は、カーナビゲーションシステム、テレビ、パソコンのインターフェースとして、使用されているが、音声による伝達手段は、ユーザが聞き逃した場合は、音声応答装置を再度、操作しないと、聞くことが出来ないため、より頑健なインターフェースを実現するために、音声を出力すると同時に、文章や単語などの文字情報を文字表示部に出力するものがある。（例えば、特許文献 1 ないし 3 参照）

また、従来の音声応答装置は、ユーザと音声で対話することにより、必要な条件を決定することでタスクを達成するものがある。

##### 【0003】

図 6 は、従来の音声応答装置を「鉄道チケット券売機」に応用した場合の、文字情報を示す図である。ユーザは、条件項目部 6 0 1 に対応した条件指定部 6 0 2 の値を、質問部 6 0 3 と音声情報とに従い埋めていくことで、所望の切符を買うことができる。

#### 【0 0 0 4】

図 7 に、従来の音声応答装置の構成図を示す。音声認識部 7 0 1 は、ユーザから音声が入力され、入力音声の認識を行うものであり、言語処理部 7 0 2 は、認識された入力音声に対して言語処理により「出発駅」や「到着駅」といったキーワードを抽出するものであり、意味理解部 7 0 3 は、ユーザの購入する切符の種類を表現するものである。

#### 【0 0 0 5】

また、対話管理部 7 0 4 は、購入する切符が一意に決まらない場合に、埋めべき条件をユーザから聞き出すような応答文に対応する内部表現を、また買うべき切符が決まった場合は、それを確認するための応答文に対応する内部表現を出力するものであり、応答文生成部 7 0 5 は、渡された応答文の内部表現から応答文を作成するものであり、応答音声出力部 7 0 6 は、スピーカなどの応答文を音声情報として出力するものであり、音声テキスト出力部 7 0 7 は、ディスプレイなどの応答文を文字情報として出力するものである。

#### 【0 0 0 6】

ここで、従来の音声応答装置では、応答テキスト出力部 7 0 7 の文字情報の出力と、応答音声出力部 7 0 6 の音声情報の出力とが同時に行われる。（例えば、特許文献 4 参照）

#### 【0 0 0 7】

##### 【特許文献 1】

特開平 1 1 - 1 4 5 9 5 5 号公報（第 9 - 1 0 頁、第 6 図）

##### 【特許文献 2】

特開平 1 1 - 3 3 9 0 5 8 号公報（第 3 - 4 頁、第 2 図）

##### 【特許文献 3】

特開 2 0 0 1 - 1 4 2 4 8 4 号公報（第 1 1 - 1 2 頁、第 3 5 図）

## 【特許文献 4】

特開平 5-216618 号公報 (第 4-5 頁、第 1 図)

## 【0008】

## 【発明が解決しようとする課題】

しかしながら、従来の音声応答装置では、ユーザの操作性向上のため、音声情報と文字情報を同時に出力していたが、ユーザは文字情報よりも、音声情報に集中するために、必ずしも期待した操作性の向上が得られないという課題があった。また、実際に、文字を見るためには、文字出力された場所に視線を動かさなければならないが、眼球運動を開始するまでには 70 msec ~ 700 msec、平均 200 msec の時間がかかることが知られている (田村博、「ヒューマンインターフェース」、オーム社、1998 年)。更に、実際に視線を文字に動かし、焦点を合わせるまでには、それ以上の時間が必要とされ、音声と文字出力を同時に出力する音声応答装置では、利用者は音声情報に集中すると考えられる。

## 【0009】

## 【課題を解決するための手段】

従来の課題を解決するために、本発明の音声応答装置は、文字情報を出力する応答テキスト出力部と、応答テキスト出力部が文字情報を出力し、遅延時間が経過した後に音声情報を出力する応答音声出力部とを含むものであり、人間の身体特性を考慮し、文字情報と音声情報の出力のタイミングを制御することができる。

## 【0010】

## 【発明の実施の形態】

本発明の請求項 1 に記載の発明は、文字情報及び画像情報を表示する応答テキスト出力部と、音声情報を音声として出力する音声出力部と、前記文字情報が出力される時刻と前記音声情報が出力される時刻とに遅延時間を設定する応答出力管理部とを含む音声応答装置としたものであり、ユーザが、文字情報を確認するまでの猶予を与えることができ、ユーザが、確実に音声情報を聞き取れることができるという作用を有する。

## 【0011】

本発明の請求項 2 に記載の発明は、遅延時間は、ユーザとの対応を行うキャラクタであるエージェントの応答テキスト出力部への表示の有無をパラメータとして決定される値である請求項 1 記載の音声応答装置としたものであり、ユーザがエージェントに注目することによって生じる眼球運動開始までの時間変化に対応できるという作用を有する。

【0012】

本発明の請求項 3 に記載の発明は、遅延時間は、応答テキスト出力部に表示される文字情報とエージェントとの距離をパラメータとして決定される値である請求項 2 記載の音声応答装置としたものであり、眼球運動に掛かる時間を考慮することができ、応答文の発声を適切なタイミングで行うことが出来るという作用を有する。

【0013】

本発明の請求項 4 に記載の発明は、遅延時間は、応答テキスト出力部に表示される文字情報の表示位置をパラメータとして決定される値である請求項 1 ないし 3 のいずれか記載の音声応答装置としたものであり、ユーザの視線移動にあわせた適切なタイミングで応答文の発声が可能であるという作用を有する。

【0014】

本発明の請求項 5 に記載の発明は、遅延時間は、ユーザ特徴を表すプロフィール情報をパラメータとして決定される値である請求項 1 ないし 4 のいずれか記載の音声応答装置としたものであり、ユーザの特徴を考慮した適切なタイミングで応答文の出力が可能であるという作用を有する。

【0015】

本発明の請求項 6 に記載の発明は、遅延時間は、操作回数、利用時間を含むユーザの利用状況をパラメータとして決定される値である請求項 1 ないし 5 のいずれか記載の音声応答装置としたものであり、ユーザの機器に対する慣れを考慮した適切なタイミングで応答文の出力が可能であるという作用を有する。

【0016】

本発明の請求項 7 に記載の発明は、ユーザの視線位置を計測する視線位置計測部を更に含み、遅延時間は、ユーザの視線位置をパラメータとして決定される値



である請求項 1 ないし 6 のいずれか記載の音声応答装置としたものであり、ユーザの視線移動にあわせた適切なタイミングで応答文の発声が可能であるという作用を有する。

【0017】

本発明の請求項 8 に記載の発明は、ユーザの呼吸の状態を測定する生体情報測定部を更に含み、遅延時間は、前記ユーザの呼吸の状態をパラメータとして決定される値である請求項 1 ないし 7 のいずれか記載の音声応答装置としたものであり、眼球運動にかかる時間をより精度よく考慮することができるという作用を有する。

【0018】

本発明の請求項 9 に記載の発明は、応答対象である文字情報は、点滅、応答対象でない文字情報よりも大きく表示又は応答対象でない文字情報とは異なる色で表示のうちのいずれか、あるいはそれらを組み合わせて表示される請求項 1 ないし 8 のいずれか記載の音声応答装置としたものであり、より確実にユーザがテキストを見ながら音声を聞くことが出来るという作用を有する。

【0019】

本発明の請求項 10 に記載の発明は、応答テキスト出力部は、文字情報を表示する場合に、ユーザの注意を喚起する報知音を出力する請求項 1 ないし 9 のいずれか記載の音声応答装置としたものであり、より確実にユーザがテキストを見ながら音声を聞くことが出来るという作用を有する。

【0020】

本発明の請求項 11 に記載の発明は、文字情報が出力される時刻と音声情報が出力される時刻とに遅延時間を設定する第 1 のステップと、文字情報を表示する第 2 のステップと、前記遅延時間の経過後に、音声情報を音声として出力する第 3 のステップとを含む音声応答方法としたものであり、ユーザが、文字情報を確認するまでの猶予を与えることができ、ユーザが、確実に音声情報を聞き取れることができるという作用を有する。

【0021】

本発明の請求項 12 に記載の発明は、文字情報が出力される時刻と音声情報が

出力される時刻とに遅延時間を設定する第1のステップと、文字情報を表示する第2のステップと、前記遅延時間の経過後に、音声情報を音声として出力する第3のステップとを含む音声応答方法を実行するプログラムとしたものであり、ユーザが、文字情報を確認するまでの猶予を与えることができ、ユーザが、確実に音声情報を聞き取れることができるという作用を有する。

#### 【0022】

以上本発明の実施の形態について、図面を参照しながら説明する。

#### 【0023】

(実施の形態1)

図1に、本発明の音声応答装置の構成を示す。音声認識部101は、ユーザから音声が入力され、入力音声の認識を行うものであり、言語処理部102は、認識された入力音声に対して言語処理により「出発駅」や「到着駅」といったキーワードを抽出するものであり、意味理解部103は、抽出したキーワードから対話の内容を判断し、対話用テンプレートを選択するものである。

#### 【0024】

また、対話管理部104は、対話用テンプレートの設定項目に抽出したキーワードを当てはめ、キーワードが不足している場合は、ユーザに不足キーワードを入力させるための応答文に対応する内部表現を出力し、キーワードが充分である場合は、対話内容を確認するための応答文に対応する内部表現を出力するものであり、応答文生成部105は、対話管理部104から入力された内部表現を用いて応答文を作成するものである。

#### 【0025】

また、応答出力管理部106は、応答文を音声情報として出力する時刻と応答文を文字情報として出力する時刻とを制御し、ユーザとの対話を行なうキャラクターであるエージェントを応答文に応じて制御するものであり、時刻情報提供部107は、現在時刻を出力するものであり、応答音声出力部108は、応答文を音声情報として出力するものであり、応答テキスト出力部109は、応答文を文字情報として出力するものである。

#### 【0026】

図2は、音声応答装置の動作を示すフローチャートである。

【0027】

まず、ステップ201にて、音声認識部101が、入力音声を認識する。次に、ステップ202にて、言語処理部102が、認識された入力音声からキーワードを抽出する。次に、ステップ203にて、意味理解部103が、キーワードから対話内容に対応したテンプレートを選択する。

【0028】

次に、ステップ204にて、テンプレートに対するキーワードが充分である場合は、ステップ205に移行し、キーワードが不足している場合は、ステップ206に移行する。次に、ステップ205にて、対話管理部104が、対話内容を確認する内部表現を作成する。次に、ステップ206にて、対話管理部104が、不足キーワードの入力を求める内部表現を作成する。

【0029】

次に、ステップ207にて、応答文生成部105が、内部表現を用いて応答文を作成する。次に、ステップ208にて、応答出力管理部106が、応答文を応答テキスト出力部109に出力し、応答文を文字情報として表示する。

【0030】

次に、ステップ209にて、応答出力管理部106が、内蔵されている時間カウンタを初期化する。次に、ステップ210にて、時間カウンタが予め定めた遅延時間  $t_h$  に達した後に、ステップ211に移行する。次に、ステップ211にて、応答出力管理部106が、応答文を応答音声出力部108に出力し、応答文を音声情報として出力する。

【0031】

以上が、本発明の音声応答装置の動作である。

【0032】

次に、応答出力管理部106による遅延時間  $t_h$  の決定方法について説明する。

【0033】

文字情報の出力と音声情報の出力との遅延時間  $t_h$  の値は、眼球運動の開始時

間と、視点移動後の焦点を合わせるタイミングを考慮して、文字情報の表示後、 $100\text{ msec} \sim 1\text{ sec}$ 程度に設定すればよい。ここで、眼球運動の開始時間は、 $70\text{ msec} \sim 700\text{ msec}$ 程度とした。

#### 【0034】

また、音声入力を行う場合に、応答テキスト出力部109に、ユーザとの対応を行うキャラクタであるエージェントを表示し、エージェントに発話させる構成とする場合は、眼球運動を開始するまでに必要とされる時間が、エージェントの表示により長くなるため、遅延時間  $t_h$  は、(数1)のように、予め定め時間を加算する方法をとればよい。なお、エージェントは、応答文生成部105が生成する応答文に合わせて、応答出力管理部106により制御される。

#### 【0035】

##### 【数1】

$$t_h = t_{h0} + \alpha 0$$

#### 【0036】

ここで  $\alpha 0$  は、エージェントが表示されている場合に、加算される調整用遅延時間であり、システムに応じて実験的に決定できる。 $t_{h0}$  は、エージェントを用いないシステム毎に、決定される遅延時間である。

#### 【0037】

ここで、応答テキスト出力部109に表示される文字情報について説明する。図3は、本実施の形態による応答テキスト出力部109に表示されるエージェント304を用いた切符購入を支援するための文字情報を示す図である。切符購入に必要とされる質問項目301として、出発駅、到着駅、禁煙と喫煙の選択、切符の枚数の質問項目301を含むテンプレートと、ユーザの回答であるキーワード302と、ユーザとの対応を行うキャラクタであるエージェント304と、ユーザに切符の購入を促す文字情報がエージェントの吹き出し303と、切符の購入を開始する開始ボタン305と、切符購入の確認を行う確認ボタン306とが表示されている。

#### 【0038】

ここで、エージェント 3 0 4 が応答テキスト出力部 1 0 9 に表示されているため、調整用遅延時間  $\alpha 0$  を用いて文字情報と音声情報の遅延時間を決定することになる。

#### 【0 0 3 9】

また、ユーザが注目するエージェント 3 0 4 と応答対象となっている文字情報との位置が離れている場合は、遅延時間  $t h$  の値を大きくする必要がある。これは、物理的に遠いものを見るためには眼球を動かす距離が長くなるためより多くの時間を必要とする事から明らかである。

#### 【0 0 4 0】

図 3 では、エージェント 3 0 4 は、応答テキスト出力部 1 0 9 の右下に配置され、出発駅の質問項目 3 0 1 は左上に配置されているため、ユーザは、最初に、出発駅の質問項目 3 0 1 を確認した後に、エージェント 3 0 4 とエージェントの吹き出し 3 0 3 とに視線を移動する。

#### 【0 0 4 1】

このため、エージェント 3 0 4 とエージェントの吹き出し 3 0 3 と質問項目 3 0 1 とを応答テキスト出力部 1 0 9 のどの位置に配置するかによって、遅延時間  $t h$  を（数 2）のように決定することができる。これは、物理的に遠いものを見るためには、眼球を動かす距離が長くなるため、より多くの時間を必要とする事から明らかである。ここで、調整用遅延時間  $\alpha 1$  は、応答テキスト出力部 1 0 9 上の位置  $p$  の関数と定義すれば良い。また、画面の中心から離れるほど調整用遅延時間  $\alpha 1$  の値を増加することで、遅延時間  $t h$  の調整を正確に行うことができる。なお、エージェント 3 0 4 が表示されていない場合は、応答テキスト出力部 1 0 9 の中心から離れるほど、調整用遅延時間  $\alpha 1$  の値が増加する方法を取ることも有効である。

#### 【0 0 4 2】

##### 【数 2】

$$t h = t h 0 + \alpha 1 (p)$$

#### 【0 0 4 3】

また、ユーザの個人差を考慮することで、詳細に遅延時間  $t_h$  を決定することも可能である。これは、眼球運動の開始のタイミング、眼球を移動速度、焦点を合わせるスピードは、加齢とともに遅くなってくると考えられ、年齢及び俊敏性によって得られるユーザの個人情報をもとに、遅延時間  $t_h$  を決定することができる。また、過去の音声応答装置の使用履歴を考慮して遅延時間  $t_h$  を決定することも可能である。

#### 【0044】

図4に、個人の年齢、俊敏性、使用履歴であるプロフィール情報が蓄積されたプロフィール情報データベース401を、応答出力管理部402が、参照する音声応答装置の構成を示す。応答出力管理部402は、プロフィール情報を参照し、遅延時間  $t_h$  を決定し、応答音声出力部108と応答テキスト出力部109とに出力する応答文の時間を調整する。

#### 【0045】

また、遅延時間  $t_h$  は、(数3)により決定することができる。ここで、 $\alpha_2$  は、ユーザの年齢をパラメータとし、年齢に伴い増加する調整用遅延時間であり、 $\alpha_3$  は、ユーザの使用履歴をパラメータとし、使用回数が多いほど減少する調整用遅延時間であり、 $\alpha_4$  は、ユーザの俊敏性をパラメータとして決定される調整用遅延時間であり、システムに応じて実験的に決定することができる。

#### 【0046】

##### 【数3】

$$t_h = t_{h0} + \alpha_2 + \alpha_3 + \alpha_4$$

#### 【0047】

また、遅延時間  $t_h$  の上限を決めておくことで、遅延時間  $t_h$  が大きくなりすぎることを避ける事ができる。なお、(数3)では、 $\alpha_2$ 、 $\alpha_3$ 、 $\alpha_4$  のすべてを用いているが、必ずしもすべてを使う必要はない。

#### 【0048】

また、切符の購入などの一連の応答動作においても、ユーザと音声応答装置との対話が進むにつれて、ユーザは音声応答装置に慣れてくると考えられる。そこ

で、このユーザの慣れ情報を利用して、遅延時間  $t_h$  の値を決めることができる。

#### 【0049】

ここで、対話の単位に応じて遅延時間  $t_h$  を変更する音声応答装置の構成を図5に示す。音声応答装置からの1つの問い合わせに対してユーザが答える応答を1つの単位として、ターン数として表す。ターン数は、対話の単位であるため、対話管理部501により求められ、応答出力管理部502に出力される。遅延時間  $t_h$  を決定する具体的な方法は、次の応答出力が  $k$  ターン目である場合は、(数4) とすることで実現できる。

#### 【0050】

##### 【数4】

$$t_h = t_{h0} + f(\alpha_5/k)$$

#### 【0051】

ここで  $\alpha_5$  は定数で、音声応答装置に応じて実験的に決めることができる。ここで  $f$  はシグモイド関数などで、値が発散しないようにする関数である。

#### 【0052】

例えば、図3に示した切符の購入では、出発駅に関する応答が第1のターンであり、切符の枚数に関する応答が第4のターンとなる。遅延時間  $t_h$  は、第1のターンより第4のターンの方が、短くなる。

#### 【0053】

また、ターン数と同様の概念として、1つの応答に対する操作時間をユーザの慣れを表す尺度として利用することが出来る。そのために、対話管理部501から得られたターン数が1であるときを、音声応答装置の使用開始と想定し、音声応答装置の使用開始時刻  $T_0$  を初期化する。その後、応答出力管理部502は応答文生成部105から応答文の入力を受けた時刻  $T$  を、時刻情報提供部107から受け取り、以下のように遅延時間  $t_h$  を、(数5) として、更新することで、ユーザの慣れに合わせて応答音声を出力することが可能となる。

#### 【0054】

## 【数5】

$$t_h = t_{h0} + f(\alpha_6 / (T - T_0))$$

## 【0055】

ここで $\alpha_6$ は定数で、音声応答装置に応じて実験的に決めることができる。ここで $T_0$ は第1ターン目の時刻、そして $T$ は現在時刻である。 $f$ はシグモイド関数などで、遅延時間 $t_h$ が発散しないようにする関数である。

## 【0056】

なお、ここでは、ターン数が1の時に音声応答装置の使用開始としたが、音声応答装置の電源を入れた時刻、開始ボタン305を押した時刻を使用開示であるとすることも可能である。

## 【0057】

また、音声出力の遅延時間は、ユーザの身体状態に合わせて行うことが有効だと考えられ、図示していないがユーザの視線を捕らえる視線測定部の測定結果を応答出力管理部106に入力し、遅延時間を求めることで、ユーザの視線位置を考慮し、視線位置のテキスト内容に合わせて、音声出力を開始すればよい。

## 【0058】

また、ユーザの呼吸に合わせて応答文を開始することにより、ユーザに心地よいタイミングで音声応答装置の応答を開始することが出来る。例えば、人間は注意を集中した瞬間は、呼吸が止まることが知られている。そこで、図示していないが、呼吸が停止した瞬間を検出する生体情報測定部の測定結果を応答出力管理部106に入力し、遅延時間を求めることで、呼吸が停止しするタイミングに合わせて応答音声を発声することができる。なお、生体情報測定部は、息を吸っているか否かは、伸縮性可変抵抗素子をつけたベルトや、呼吸に伴う胸囲や腹部の変化を測ることで実現できる（田村博、「ヒューマンインターフェース」、オーム社、1998年）。

## 【0059】

以上、本発明の実施の形態において、説明した遅延時間の決定方法は、それらを組み合わせることで、より詳細に制御できるものである。



**【0060】**

また、テキストの表示方法を工夫することで、より確実にユーザにテキスト情報に注目させることが可能である。例えば、テキストの表示の点滅、表示した瞬間は文字を予め定めた大きさよりも大きく表示し、徐々に予め定めたの大きさに文字に変化させる方法でユーザの注意を引くことが可能である。また、文を構成する単語のうち、名詞、動詞、形容詞といった機能語の文字の大きさを大きくしておく、さらには色を変えておくといったことで、よりユーザに分りやすい表示方法が可能となり、注意を引きやすくなる。また、テキストを表示する瞬間に、報知音を発生させることも有効である。

**【0061】**

なお、音声応答装置は、テレビ、カーナビゲーションシステム、携帯電話、携帯端末、パーソナルコンピュータ、電話機、ファックス、電子レンジ、冷蔵庫、掃除機などテキスト情報と、音声情報を出力可能な機器であればどのような機器に対しても本発明による応答制御を使用することが出来る。なお、ターン数は、インターフェースの操作回数と置き換えられる。

**【0062】**

以上の構成によれば、応答出力管理部106が、音声情報とテキスト情報との出力時間に遅延時間を設けることで、ユーザは、まず、テキスト情報に集中することが可能となり、音声情報を聞き逃すことなく、音声応答装置を利用することができる。

**【0063】****【発明の効果】**

人間の認知特性を考慮し、文字出力と音声出力のタイミングを制御することで、応答文を聞き取り易くする効果がある。

**【図面の簡単な説明】****【図1】**

本発明の第1の実施の形態による音声対話装置を示す図

**【図2】**

本発明の第1の実施の形態による応答出力管理部のフローチャート

## 【図 3】

本発明の第 1 の実施の形態による文字情報を示す図

## 【図 4】

本発明の第 1 の実施の形態による音声対話装置を示す図

## 【図 5】

本発明の第 1 の実施の形態による音声対話装置を示す図

## 【図 6】

従来 of 文字情報を示す図

## 【図 7】

従来 of 音声対話装置を示す図

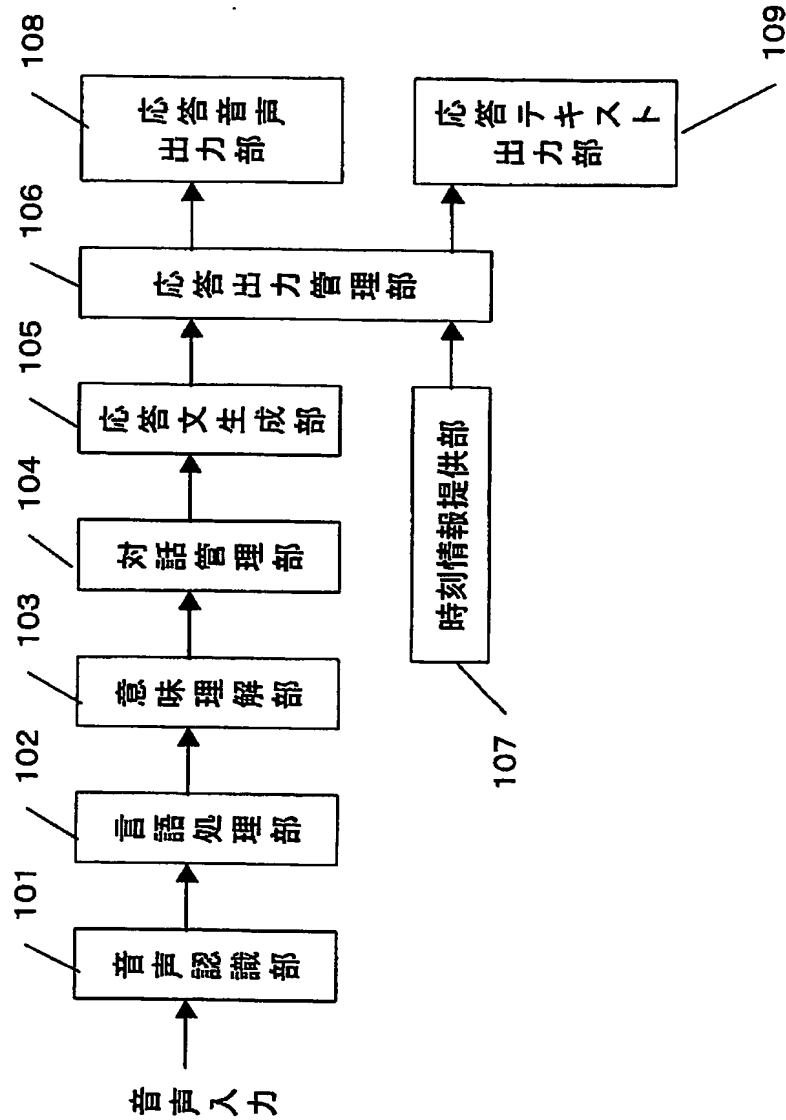
## 【符号の説明】

- 1 0 1 音声認識部
- 1 0 2 言語処理部
- 1 0 3 意味理解部
- 1 0 4、5 0 1 対話管理部
- 1 0 5 応答文生成部
- 1 0 6、4 0 2、5 0 2 応答出力管理部
- 1 0 7 時刻情報提供部
- 1 0 8 応答音声出力部
- 1 0 9 応答テキスト出力部
- 3 0 1 質問項目
- 3 0 2 キーワード
- 3 0 3 吹き出し
- 3 0 4 エージェント
- 3 0 5 開始ボタン
- 3 0 6 確認ボタン
- 4 0 1 プロファイル情報データベース

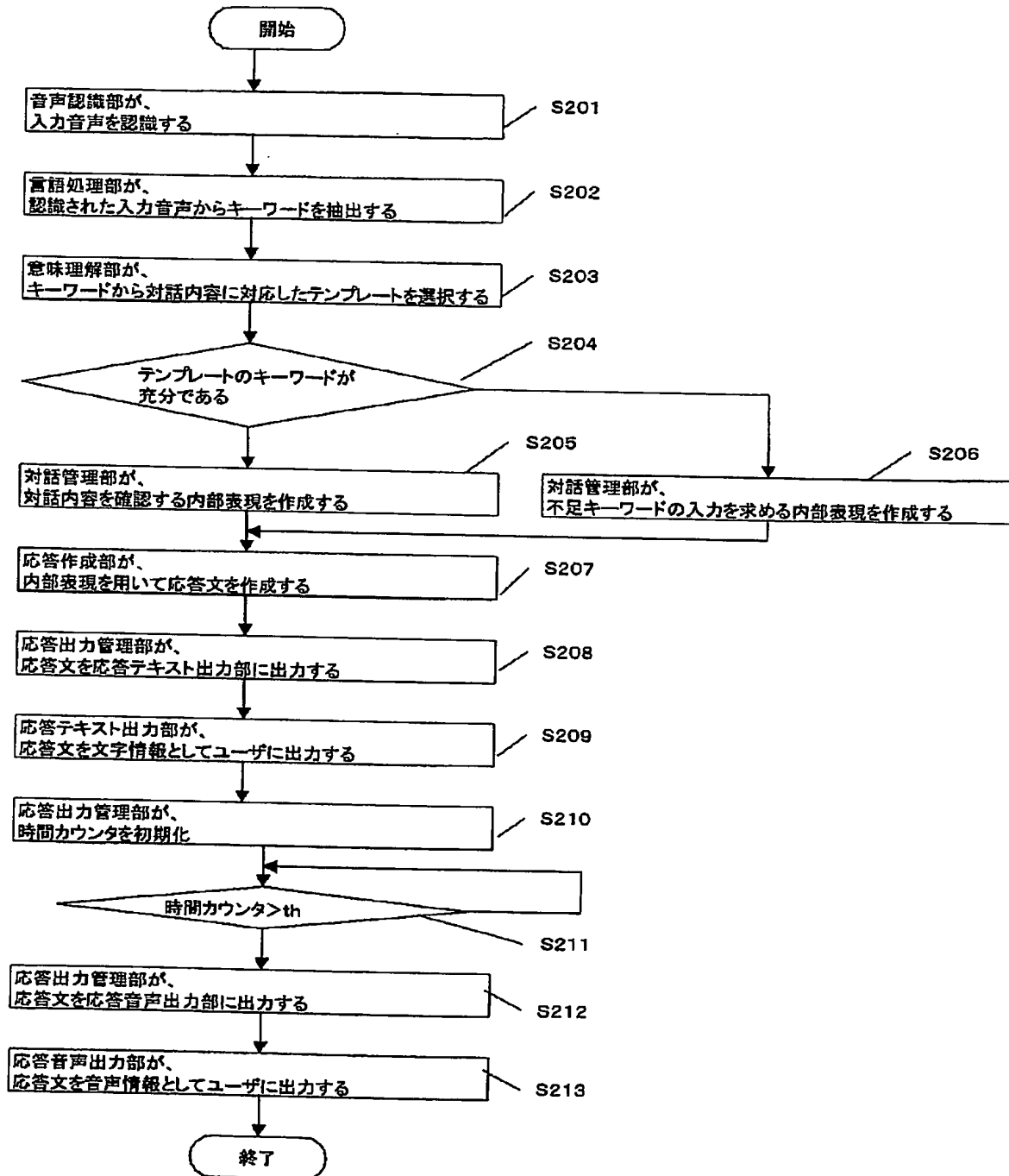
【書類名】

図面

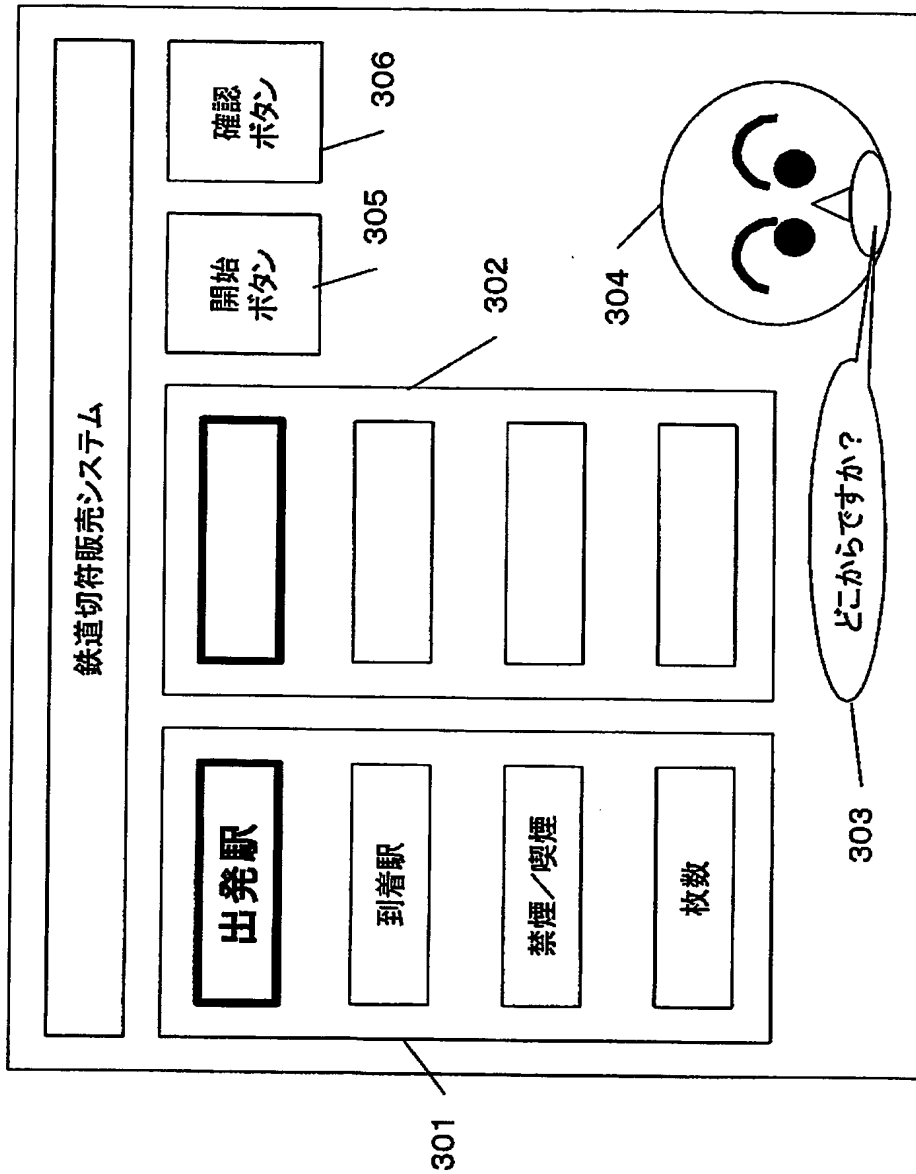
【図 1】



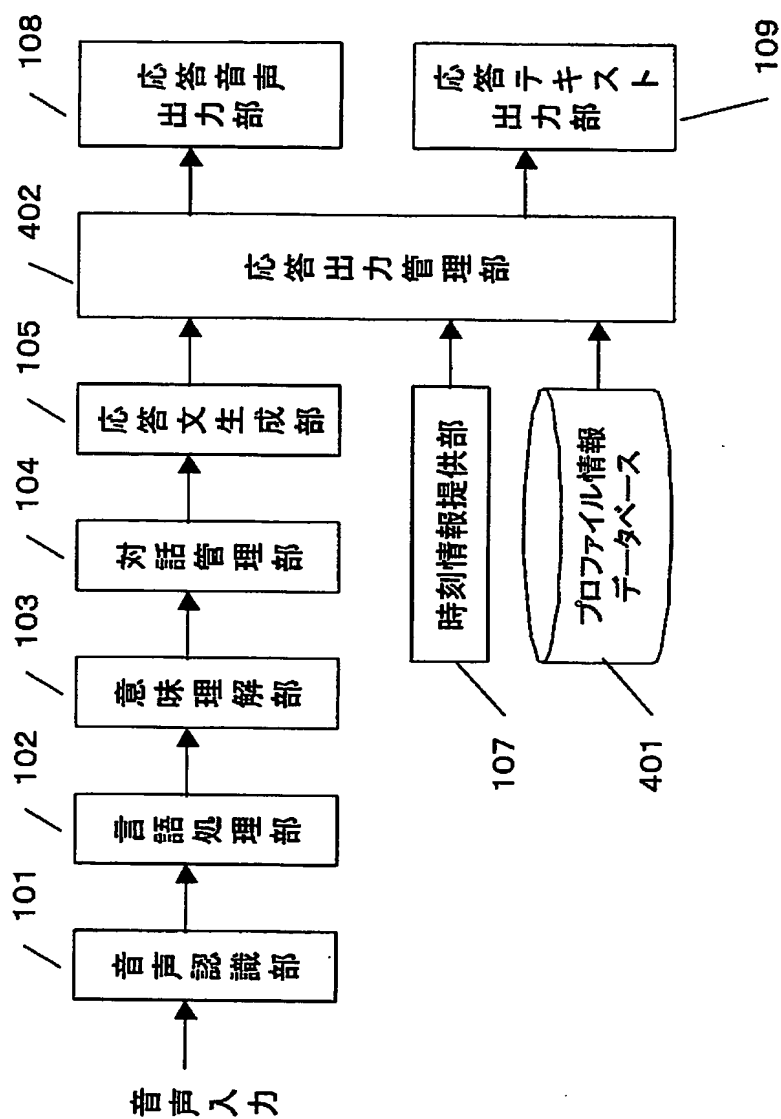
【図 2】



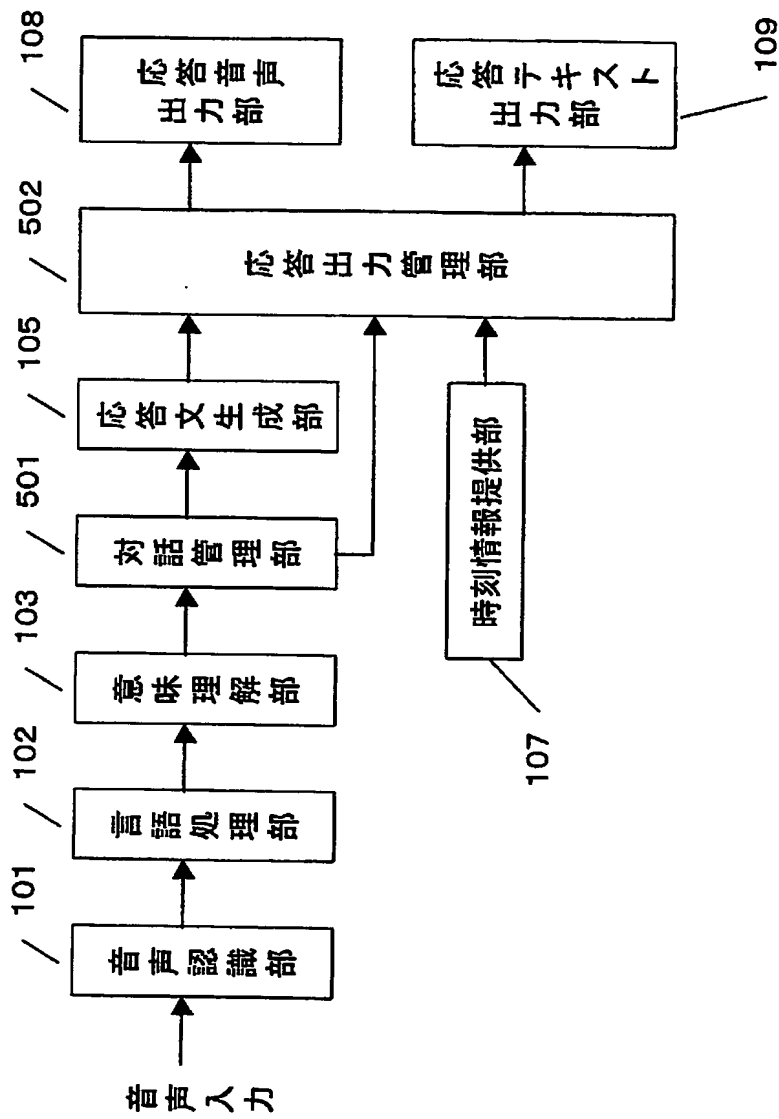
【図 3】



【図 4】



【図 5】



【図 6】

鉄道切符販売システム

601 出発駅

到着駅

禁煙／喫煙

枚数

602 A駅

B駅

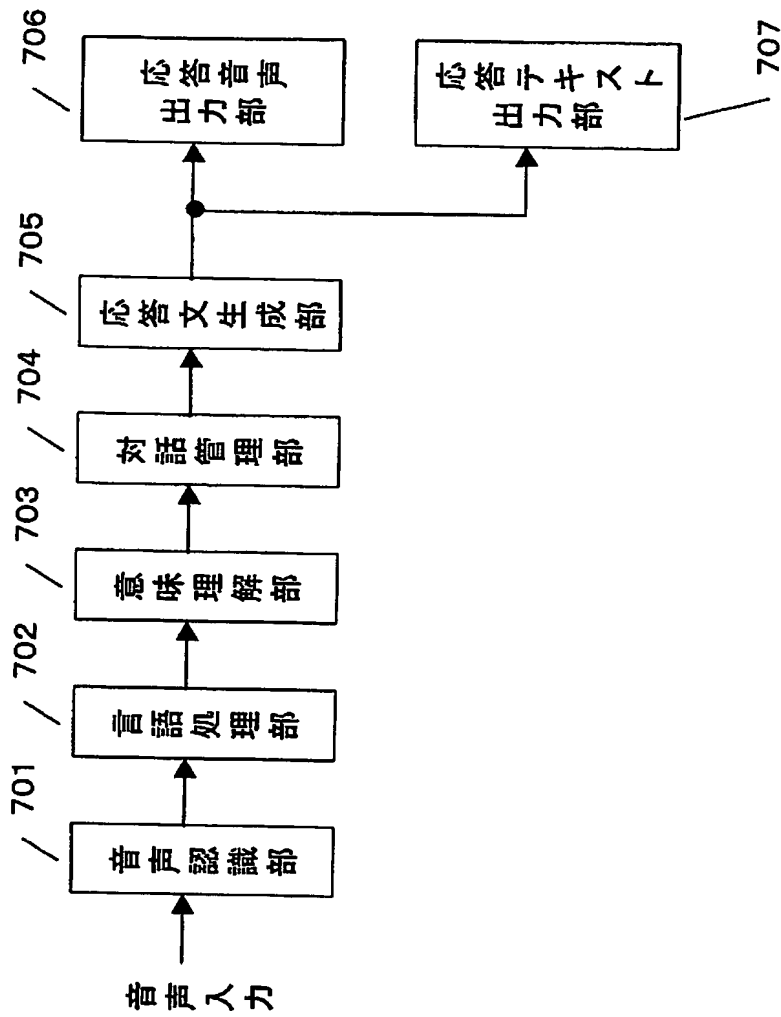
禁煙

3枚

603 以上よろしいでしょうか?



【図 7】



【書類名】 要約書

【要約】

【課題】 ユーザの操作性向上のため、音声情報と文字情報を同時に出力していたが、ユーザは文字情報よりも、音声情報に集中するために、必ずしも期待した操作性の向上が得られないという課題があった。

【解決手段】 文字情報を出力する応答テキスト出力部109と、応答テキスト出力部109が文字情報を出力し、遅延時間が経過した後に音声情報を出力する応答音声出力部108とを含むものであり、人間の身体特性を考慮し、文字情報と音声情報の出力のタイミングを制御することができる。

【選択図】 図1

出願人履歴情報

識別番号

[000005821]

1. 変更年月日

1990年 8月28日

[変更理由]

新規登録

住所

大阪府門真市大字門真1006番地

氏名

松下電器産業株式会社